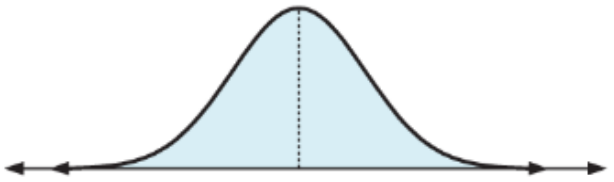
## Warm Up # 4-3

1. Given: X ~ N(4.2, 0.36) and P(X > k) = 0.87

   Illustrate, show what you enter in your calculator and find k:



## HW Questions: p. 309

**EXERCISE 10C**

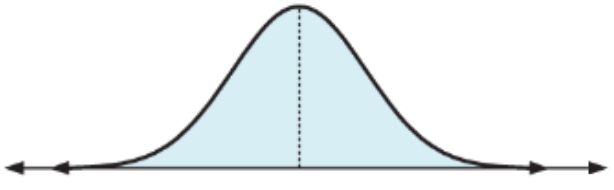
**1** Suppose $X \sim \mathrm{N}(20, 3^2)$. Illustrate with a sketch and find $k$ such that:

**a** $\mathrm{P}(X \leqslant k) = 0.348$ **b** $\mathrm{P}(X \leqslant k) = 0.878$ **c** $\mathrm{P}(X \leqslant k) = 0.5$

**2** Suppose $X \sim \mathrm{N}(38.7, 8.2^2)$. Illustrate with a sketch and find $k$ such that:

**a** $\mathrm{P}(X \leqslant k) = 0.9$ **b** $\mathrm{P}(X \geqslant k) = 0.8$

$P(X \leq k) = 0.2$



**3** Suppose $X \sim \mathrm{N}(30, 5^2)$ and $\mathrm{P}(X \leqslant a) = 0.57$.

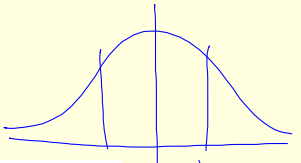
**a** Using a diagram, determine whether $a$ is greater or less than 30.

**b** Use technology to find $a$.

**c** Without using technology, find: **i** $\mathrm{P}(X \geqslant a)$ **ii** $\mathrm{P}(30 \leqslant X \leqslant a)$

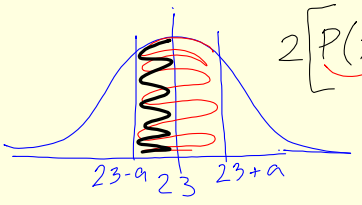
**4** Given that $X \sim \mathrm{N}(23, 5^2)$, find $a$ such that:

**a** $\mathrm{P}(X < a) = 0.378$ **b** $\mathrm{P}(X \geqslant a) = 0.592$



**c** $\mathrm{P}(23 - a < X < 23 + a) = 0.427$

68%




$2[P(x<23) - P(x<23-a)] = 0.427$

$0.5 - P(x<23-a) = 0.2135$

$P(x<23-a) = 0.2865$

$23 - a = \text{invNorm}(0.2865, 23, 5)$

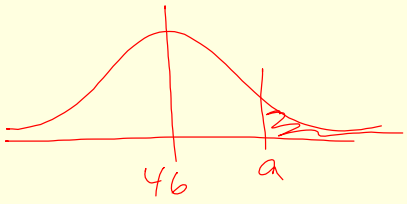
$a = 23 - \text{invNorm}(0.2865, 23, 5)$

$a \approx$

**5** The students of Class X sat a Physics test. The average score was 46 with a standard deviation of 25. The teacher decided to award an A to the top 7% of the students in the class. Assuming that the scores were normally distributed, find the lowest score that would achieve an A.

$X \sim N(46, 25^2)$

$P(X < a) = 0.93$




**6** The lengths of a fish species are normally distributed with mean 35 cm and standard deviation 8 cm. The fisheries department has decided that the smallest 10% of the fish are not to be harvested. What is the size of the smallest fish that can be harvested?

**7** The lengths of screws produced by a machine are normally distributed with mean 75 mm and standard deviation 0.1 mm. If a screw is too long it is automatically rejected. If 1% of screws are rejected, what is the length of the smallest screw to be rejected?

**8** The weights of cabbages sold at a market are normally distributed with mean 1.6 kg and standard deviation 0.3 kg.

**a** One wholesaler buys the heaviest 10% of cabbages. What is the minimum weight cabbage he buys?

**b** Another buyer chooses cabbages with weights in the lower quartile. What is the heaviest cabbage this person buys?

**9** The volumes of cool drink in bottles filled by a machine are normally distributed with mean 503 mL and standard deviation 0.5 mL. 1% of the bottles are rejected because they are underfilled, and 2% are rejected because they are overfilled; otherwise they are kept for retail. What range of volumes is in the bottles that are kept?

# Statistical Applications

(a 9-10 day unit)

Today: Correlation and LSRL

Handouts are coming your way, so for now, just listen...

Often a statistician will want to know how often two variables are "related" or "associated".

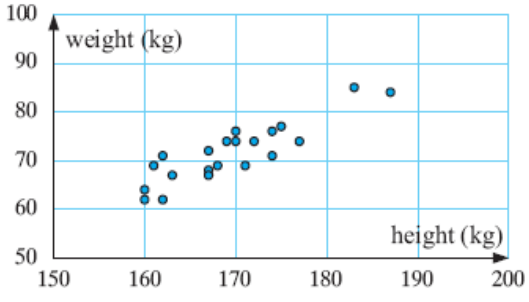
examples

outside temperature vs. # customers at Starbucks

incidence of heart disease vs. intake of Omega 3 oils

arm length vs. running speed

**Individual data is often displayed as a Scatter Diagram or Scatter Plot**



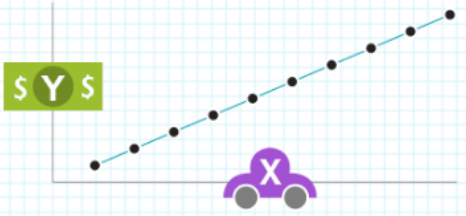

**The points are not joined together.**

**The resulting distribution of dots can then *give us an idea of whether the relationship between the two variables is weak or strong or can be modeled by a function.***

# Correlation Basics:

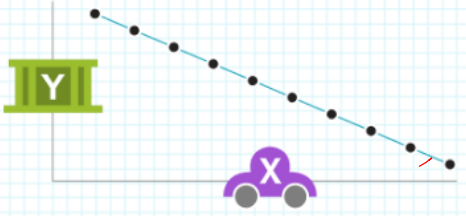
**Example 1:** Let's say **X** is the number of cars you buy, and **Y** is the amount of money you spend (assuming all the cars cost the same). The more cars you buy, the more money you spend.

**We call this a Positive Correlation.**




**Example 2:** Let's say **X** is still cars, but now **Y** is your bank account balance. With each car you buy, your bank account gets smaller and smaller. As **X** goes up, **Y** goes down.

**We call this a Negative Correlation.**

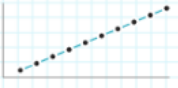



Linear

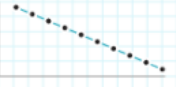
Linear

The **Correlation Coefficient (r)** measures the strength of the correlation.

Values for **(r)** range from **+1.0**

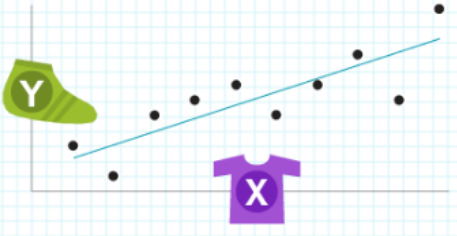


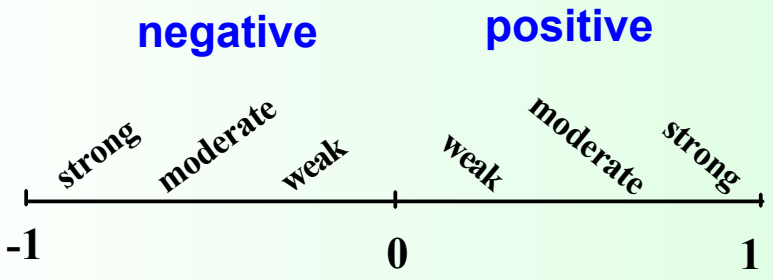
to **-1.0**



for r = ±1, the relationship between the variables is perfectly linear.

**Example 3:** Let's try a less idealized example. Here, **X** is shirt size, and **Y** is shoe size. As one size goes up, so does the other, but the relationship varies from person to person. The correlation here is still positive, but it's not perfect (**r = +0.83**).






# Values of *r*

## Important Note:

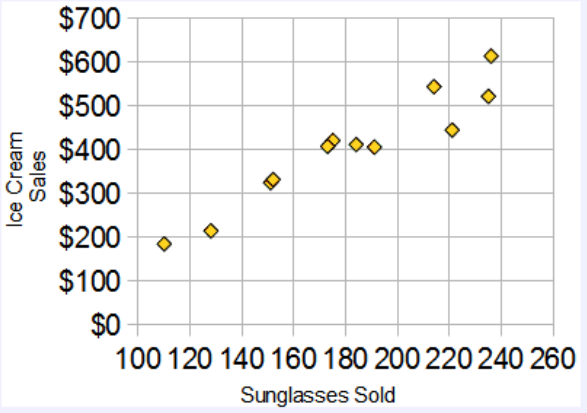
**The existence of correlation does not, repeat NOT, imply that one variable is CAUSED by the other.**

**It simply shows that the two variables are related.**

**Both of them could be caused by a third variable you don't even know about.**

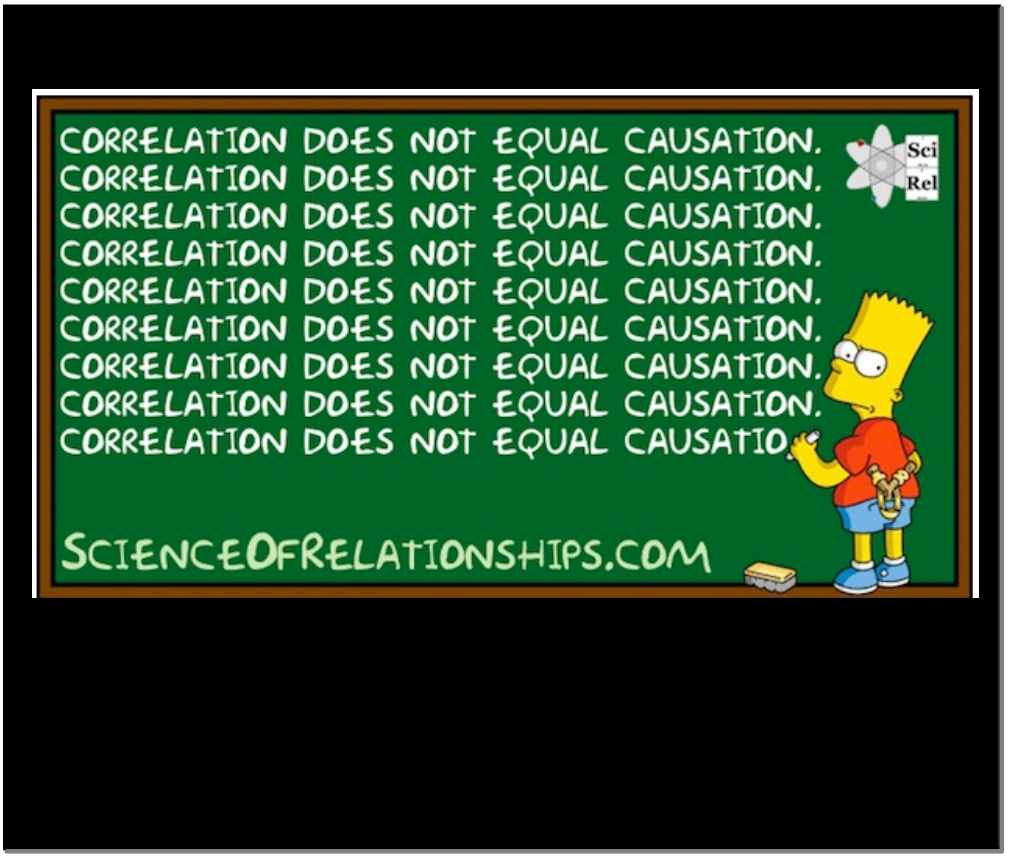
## Example: Sunglasses vs Ice Cream

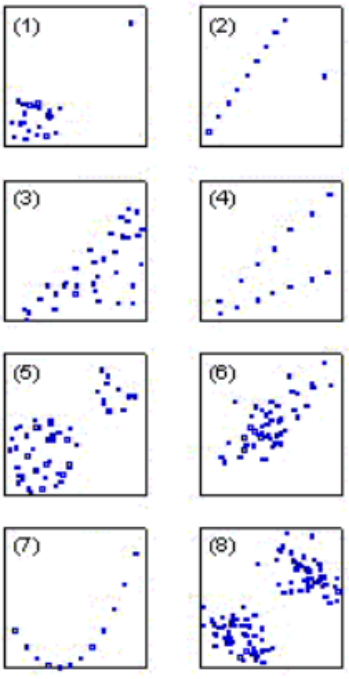
Our Ice Cream shop finds how many sunglasses were sold by a big store for each day and compares them to their ice cream sales:




The correlation between Sunglasses and Ice Cream sales is high

Does this mean that sunglasses make people want ice cream?






"To further illustrate the problems of attempting to interpret a correlation coefficient without looking at the corresponding scatterplot, consider this set of scatterplots, which duplicates most of the examples from pages 78-79 of *Graphical Methods for Data Analysis* by Chambers, Cleveland, Kleiner, and Tukey. " (*).

**"Each data set has a correlation coefficient of 0.7."**

"The moral of these displays is clear: **ALWAYS LOOK AT THE SCATTERPLOTS!** "

"The correlation coefficient is a numerical summary and, as such, it can be reported as a measure of association for any batch of numbers, no matter how they are obtained. Like any other statistic, its proper interpretation hinges on the sampling scheme used to generate the data. "

Always start a study with a scatter plot

# Hand out on the Correlation Coefficient, r

* Read through this handout for a deeper understanding.

Note: This scale on the back of your handout is not absolute! There are different scales like this published and they don't all agree!

r = -1 indicates a **perfectly strong** *negative* linear relationship.

r = -0.8 indicates a **relatively strong** *negative* linear relationship

r = -0.5 indicates a **moderate** *negative* linear relationship

r = -0.2 indicates a **weak** *negative* linear relationship

r = 0 indicates **no** *linear* relationship

r = 0.5 indicates a **moderate** *positive* linear relationship

r = 0.8 indicates a **relatively strong** *positive* linear relationship

r = 1 indicates a **perfectly strong** *positive* linear relationship.

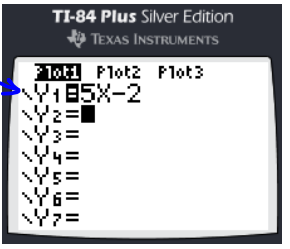
*On the back of the next handout:

## Conflicts

Scatter plots and other Stat Plots can cause trouble when you graph functions in the "Y=" menu.

Therefore, turn **off** Stat plots when you are done.

You can also see them turned on in the Y= menu

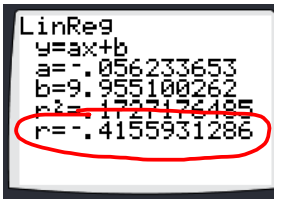



## Linear Correlation Coefficient, r

Same steps as LSRL

Notice the correlation coefficient, *r*, is given on the last line.

**If** you don't see it, then you need to turn your "DiagnosticsOn" in the Catalog.



**Correlation Classwork** Name _____ Team_____ Per_____

**Part 1 - Can one's success in Algebra 1 predict success in Geometry?**

| Students | Algebra | Geometry |
|---|---|---|
| A | 34 | 32 |
| B | 67 | 39 |
| C | 56 | 59 |
| D | 78 | 71 |
| E | 44 | 70 |
| F | 82 | 79 |
| G | 91 | 85 |

- ✓ Plot the data shown on a scatter plot with "Algebra" as the independent variable (x-axis). LABEL each axis.
- ✓ Calculate the mean Algebra score, $\bar{x} \approx$ _____ and the mean Geometry score, $\bar{y} \approx$ _____
- ✓ Plot the mean point, $(\bar{x}, \bar{y})$, on the graph and mark with "M".
- ✓ Calculate the Pearson's Correlation Coefficient, $r \approx$ _____

... There's more.

WS handed out in class.

Pick up all three handouts if you are absent.

## Viewing Data in a Scatter Plot

Two variable data should already be entered

Turn on your Statistical Plots:
"**STAT PLOT**"


Select Stat Plot 1 --then you should see this


Toggle to **On**, select scatter plot, Match your lists to the location of your data.

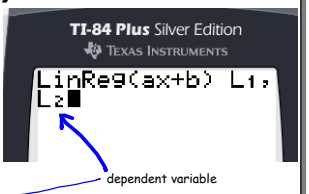
Then, **ZOOM 9**, (which is **ZoomStat**) and then you should see your Scatter plot of your data.

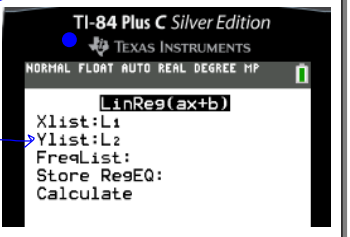
When finished, be sure to turn OFF the STAT PLOT.

# Calculating the Line of Best Fit (LSRL)

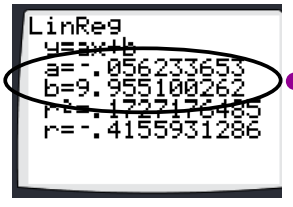
Select **STAT** then toggle to **CALC**, then to **LinReg(ax + b)**
Then the two lists which contain your data with a comma in between.

Select **ENTER** and the top two lines will give your *slope* and *y-intercept*




TI-83 or TI-84 TI-84 Plus

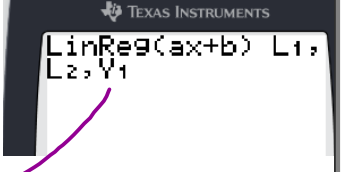



$y = -.056x + 9.96$

# Viewing your line and your Scatter Plot simultaneously

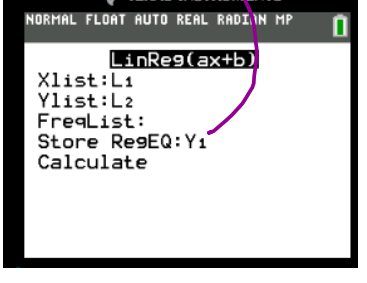
Select **STAT** then toggle to **CALC**, then to **LinReg(ax + b),** Then add one more comma and $Y_1$ , then **ENTER**. Then **ZOOM 9**

TI-84 Plus




to find $Y_1$ look for **VARS** then toggle to **Y-VARS** then **Function**

TI-84 PLUS C

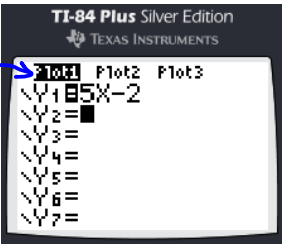


## Conflicts

Scatter plots and other Stat Plots can cause trouble when you graph functions in the "Y=" menu.

Therefore, turn **off** Stat plots when you are done.

You can also see them turned on in the Y= menu

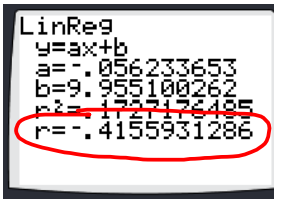



## Linear Correlation Coefficient, r

Same steps as LSRL

Notice the correlation coefficient, *r*, is given on the last line.

**If** you don't see it, then you need to turn your "DiagnosticsOn" in the Catalog.




HW: 11A,

p. 319 # 1, 2, 4, 5

HW: 11A,

p. 319 # 1, 2, 4, 5

correlation viewing

**http://wilderdom.com/301/int/cor-guess.html**